21531

RECEIVED
MAR 15 1974

Network Working Group
Request for Comments: 615

D. Crocker (UCLA-NMC)
1 MAR 74

NIC #21531

Postel

Proposed Network Standard Data Pathname Syntax

There seems to be an increasing call for a Network Standard Data Pathname (NSDP); that is, a standardized means of referring to a specific location for/of a collection of bits somewhere on the Network.

The reasons for a standard or virtual anything have been discussed, at length, elsewhere and will not be elaborated upon here. Rather than attack the entire issue of virtual pathnames, I wish only to propose a standardized SYNTAX for specifying pathnames. Such a standard will be useful for 1) users who are unfamiliar with a site or who use several different sites and do not want to have to remember each site's idiosynchracies, 2) programs accessing data at several other sites, and 3) documentation.

The syntax allows the user to specify the necessary network, host, peripheral device, directory, file, type, and site-specific fields. Adding other fields, as needed, is expected to be quite simple.

First the BNF:

```
<NSDP>      ::=  % <bulk> <cr><lf>

<bulk>      ::=  <field> / <field> <bulk>

<field>     ::=  <key> <L-delim> <name> <R-delim>

<key>       ::=  NETWORK / HOST / PERIPHERAL/ DIRECTORY /
                 FILE / TYPE / SITEPARM / N / H / P / D / F /
                 T / S
```

<L-delim> ::= any printable character that is not in the succeeding <name> field and that is acceptable to the object site. For visual aesthetics and to facilitate human parsing, anytime <L-delim> is a left-bracket character (<, [, (, ¬), <R-delim> must be the complementary right-bracket character (>,],), |). 4b5

<name> ::= any sequence of characters acceptible to the object site. This is the actual data field with the file, directory, device (or whatever) name. 4b6

<R-delim> ::= Either 1) the same character as <L-delim> or 2) if the <L-delim> character is a left-bracket character (<, [, (, ¬) then its complementary right-bracket (>,],), |). 4b7

<cr> ::= carriage-return 4b8

<lf> ::= line-feed 4b9

And some elaboration: 4c

The syntax allows <name> fields to be an arbitrary number of characters long. Case is irrelevant to the syntax, though some sites will care about case in <name> fields. 4c1

<Key> indicates what part of the pathname the next <name> is going to refer to. The single-character keys are abbreviations for the respective full-word keys. 4c2

<Fields> ARE order dependent, but defaulted ones may be omitted. The order is as indicated for <key>s. That is, Network, Host, ... Siteparm. 4c3

Fields may be repeated, as appropriate for the object site; that is, multiple Directory fields, etc. 4c3a

The validity of any combination of <field>s is entirely site-dependent. For example, if a site will accept it, an NSDP with a Host field, and nothing more, is permissible. 4c3b

<delim> is used to delimit the beginning and end of the <name> field. 4c4

Explanation of <key>s: 4d

NETWORK or N: Currently, only ARPA is defined. 4d1

HOST or H: Reference to host, by official name or nickname or number. The default radix is ten; a numeric string ending with "H" indicates hexadecimal, "O"(oh) indicates octal, and (gratuitously) "D" indicates decimal. 4d2

PERIPHERAL or P: Peripheral device being referred to. 4d3

DIRECTORY or D: Name of a directory which contains a pointer to the entity (directory or filename) specified in the following <field>. 4d4

FILE or F: Basic name of the file or data set. 4d5

TYPE or T: Optional modifier to filename. (Tenex calls it the extension.) 4d6

SITEPARM or S: A parameter, such as an access specification or version number, peculiar to the object site. The content of the <name> field must serve to identify what Siteparm is involved. Each site will be responsible for defining the syntax of Siteparm <name>s it will accept. 4d7

Some reserved PERIPHERAL <name>s: 4e

DISK or DSK: Immediately accessible, direct-access storage. 4e1

ONLINE or ONL: Whatever immediately-accessible (measured in fractions of a second) storage the user accesses by default; usually disk. 4e2

TAPE or TAP: Industry-compatible magnetic tape. 4e3

TAPE7 or TP7: 7-Track industry compatible tape. 4e4

TAPE9 or TP9: 9-Track industry compatible tape. 4e5

DECTAPE or DEC: DEC Tape. 4e6

OFFLINE or OFF:	Any tertiary storage; usually tape, though "devices" like the Datacomputer are permissible. The user should expect to wait minutes or hours before being able to access OFFLINE files.	4e7
PRINTER or PTR:	Any available line-printer.	4e8
DOCPRINTER or DOC:	Upper-lower case line printer, preferably with 8 1/2" X 11" unlined paper.	4e9
PAPER or PAP:	Paper tape.	4e10
PUNCH or PUN:	Standard 80-column card punch.	4e11
READER or RDR:	Standard 80-column card reader.	4e12
OPERATOR or OPR:	System Operator's console.	4e13
CONSULTANT or CON:	On-line consultant.	4e14

Defaults: 4f

Defaults will generally be context dependent. Consequently, the following defaults are offered only as guidelines: 4f1

Network:	ARPA	4f2
Host:	The host interpreting the NVP	4f3
Peripheral:	ONLINE (DISK)	4f4
Directory:	The user's current "working" directory, usually set by the logon process.	4f5
Filename:	None.	4f6
Type:	None.	4f7
Siteparm:	None.	4f8

General Comments 5

The only field that must be considered in relation to any host's current syntax is the escape-to-NVP field (The per-cent sign as the first character of a pathname specification). It is not currently known to conflict with any host's syntax. 5a

Exclamation mark (!) is the only other character that seems permissible (on the assumption that the character should be a graphic). Its use would cause minor problems at Multics; but more importantly as a graphic, it is too similar to the numeral "1". 5a1

The syntax is intended to be adequate for all hosts, so any given portion of it may be inappropriate for any given host. 5b

A site is expected to permit specifications in a given field iff that site already has a way of accepting the same information. 5b1

I believe that any modifications to the syntax will be graceful additions, rather than wholesale redesign, and thus can be deferred for a while. Currently, any undefined attributes must be specified in a Siteparm field. 5c

Perhaps Version, Access protection and Accounting, as well as other types of information, should be made standard <key>s, rather than buried as Siteparms. I expect that the next version of the NSDP Syntax specification will include them as <key>s, but I would like to wait for some comments from the community. 5c1

The syntax does not currently allow addressing any collection of bits samller than a file. This can be remedied by adding PAGE, BYTE and other <key>s; but, again, I would like to solicit some comments first. 5c2

Disclaimer 6

A pathname specified in the proposed syntax is fairly easy to type but is quite ugly to read. So, at the expense of design cleanliness, the <L-delim>/<R-delim> syntax was modified in an attempt to remedy the problem somewhat. As you will see below, it is only partially successful. 6a

The first draft of this document nad a syntax that was a mix of Tenex and Multics conventions. That is, 6b

(Network)/Host/Peripheral:Directory>Filename.Type;Siteparm 6b1

Though visually more attractive and generally quicker to type, it lacks extensibility. For example, adding Version number or Access protection as standard fields would be difficult. 6c

It is suggested that human interfaces be built to translate to/from NSDP syntax and the user's standard environment. 6d

Some sample pathnames: 7

%H/ISI/D<DCROCKER>F(MESSAGE)T/TXT/S(P770404)<cr><lf> refers to my protected message file at ISI (<DCROCKER>MESSAGE.TXT;P770404). 7a

%H/OFFICE-1/D>JOURNAL>F<18659>T.NLS.<cr><lf> refers to NIC Journal document #18659 (Tenex file <JOURNAL>18659.NLS). 7b

%H/65/D.ARPO61.D.LAD.F.DOCUMENT.<cr><lf> refers to a file ARPO61.LAD.DOCUMENT at UCLA-CCN. Note the use of multiple Directory fields. 7c

%H/540/D//D>udd>D>Comp=net>D>Map>F(Mail)<cr><lf> refers to file >udd>CompNet>Map>Mail at Mit-Multics. Note that the initial NSPD Directory <name> field is empty. This conforms to Multics' method of starting at the top of its directory structure. 7d

I would like to thank Jon Postel, Vint Cerf, Jim White, Charlie Kline, Ken Pogran, Jerry Burchfiel and Tom Boynton for their suggestions. 8